LLMs and Politics

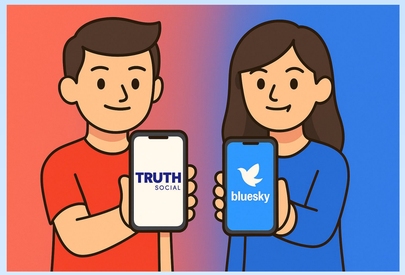
Northeastern University
Khoury College of Computer Sciences

Xingrui Huang, Liangyou Xu
Khoury College of Computer Science, Northeastern University

Supervised by:
Prof Tony Mullen, Yifan Hu

1. Background

- Online political discourse is increasingly polarized across ideologically distinct platforms.
- Existing stance/bias models are trained on news/Twitter and may not generalize to Truth Social or Bluesky.
- Research questions: How do cross-platform stance models behave, and what biases do they introduce?



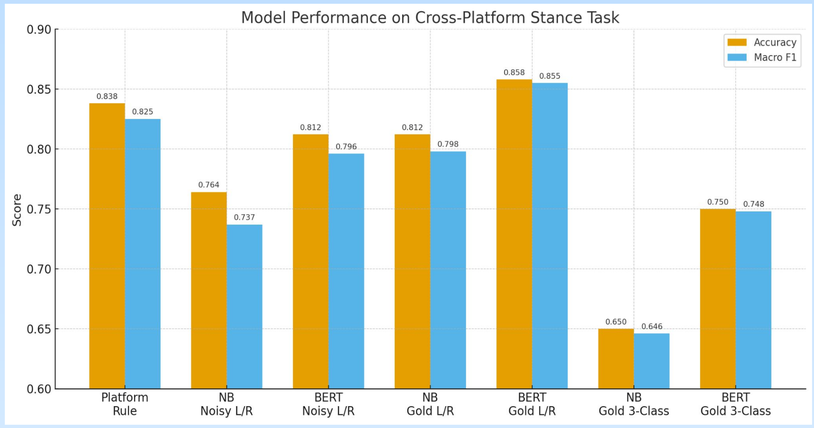
2. Datasets

- Curated Bluesky & Truth Social Posts
  - 80k English posts from Bluesky and Truth Social over 20 political topics
- Early-Platform Comparative Corpus (for clustering)
  - Bluesky (February 2023 to March 2024)
  - Truth Social (February 2022 to October 2022)
- US Congress tweets (2008–2017)
- Politics.com Forum posts (2005)

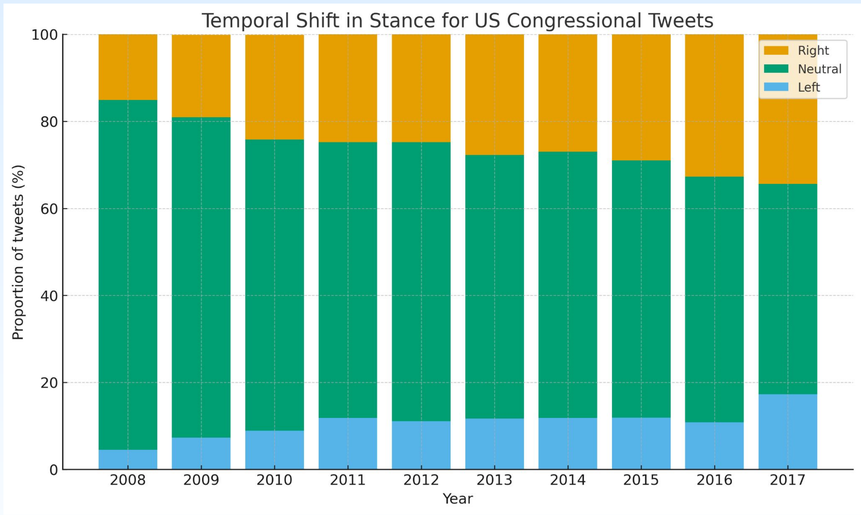
3. Methodology

- Labeling: Noisy binary labels from platform (Bluesky ≈ Left, Truth Social ≈ Right); LLM-generated labels used as Gold Left / Neutral / Right.
- Models: Complement Naive Bayes (TF–IDF 1–2-grams) and DistilBERT fine-tuning on both noisy and Gold labels (binary L/R and 3-class L/N/R).
- Tasks: In-domain stance prediction, cross-platform transfer, and temporal analysis on Congressional tweets and political forum posts.

4. Experimental Results

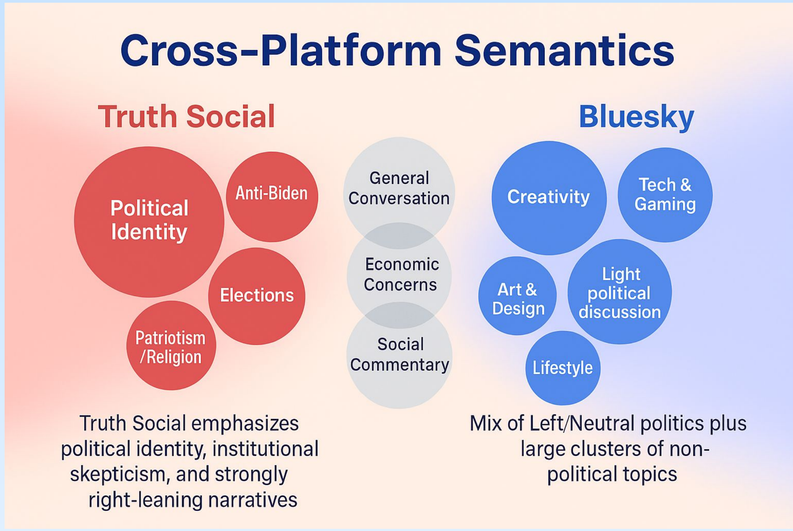


- DistilBERT substantially outperforms Naive Bayes on 3-class stance (macro-F1).
- Binary Left/Right BERT model outperforms a simple platform-only baseline.
- Adding Neutral makes the task harder; contextual embeddings help with ambiguous posts.



Temporal Polarization in Congress Tweets

- Apply the BERT L/N/R model to US Congressional tweets (2008–2017).
- Neutral share drops; explicit Left/Right stance rises over time.
- Reveals growing partisan expression in official communications under a modern stance model.



Cross-Platform Semantics

- Truth Social: overwhelmingly Right-labeled, emotionally polarized
- content.
- Bluesky: mix of Left/Neutral politics plus large clusters of non-political topics.
- Shared topics exist, but each platform emphasizes distinct narratives and styles.

5. Conclusion & Future Work

- Cross-platform LLM-assisted labeling produces a reusable stance dataset.
- DistilBERT uncovers polarization trends but also imports platform-specific biases.
- Next: human validation, causal analysis of cross-domain bias, and extension beyond US politics.